

修 士 論 文 の 和 文 要 旨

大学院		電気通信学	研究科	博士前期課程	情報工学	専攻
氏	名	伊與田 達也			学籍番号 0431007	
論文題目		Web写真ニュースの分類と検索				
<p>要 旨</p> <p>現在World Wide Web(以下、Web)上には情報が氾濫しており、毎日大量のニュース記事が配信されている。それらを毎日自動的に収集することによって、個人でも容易にニュースデータベースを構築することが可能である。しかしながら一ヶ月間で1000件以上にもなるニュース記事からユーザが興味のある記事を探し出すことは容易ではない。そこで、大量のニュース記事からユーザが興味のある記事を提示するシステムが望まれる。</p> <p>本研究では、Web上にある写真付きのニュースを対象にして、ニュース記事のテキスト情報に加えて、ニュース写真の画像特徴量も用いて、ジャンル(スポーツ、コンピュータなど)や事件、トピック毎にクラスタリングするシステムを提案する。写真付きニュースは、写真によって直観的に記事の内容を把握できる特徴があり、それらをクラスタリング表示することによって、容易に興味あるニュースを見付け出すことが可能となる。また、時系列によるクラスタリングを用いて、事件や出来事の経過を表示することによって、普段ニュースを見ていないユーザが、一目見るだけである期間の社会の出来事がわかるようなシステムの作成を目的とする。</p> <p>本研究では、写真によって直観的に記事の内容を把握できるという特徴がある写真付きニュースを利用して、記事のテキスト情報と写真の画像情報の両方を用いて標準のクラスタリング(k-means)を行い、大量のニュースの中からユーザにとって興味あるニュースを容易に見付け出すことが可能なシステムを提案・実装をし、検証を行った。また、時系列によるクラスタリングにより、事件や出来事の流れを表すスレッドを作成して、視覚的な表示を行うことにより、ニュース記事のつながりを直観的に理解できるシステムを提案、実装した。評価実験を行った結果、標準のクラスタリングが記事羅列表示よりも大量の記事から興味ある記事を見つけやすいことが確認できた。</p>						